

## **RAW SEQUENCE LISTING**

**The Biotechnology Systems Branch of the Scientific and Technical  
Information Center (STIC) no errors detected.**

Application Serial Number: 08/894,3560  
Source: 1Fw/b  
Date Processed by STIC: 10/11/05

# ***ENTERED***



IFW16

## RAW SEQUENCE LISTING

DATE: 10/11/2005

PATENT APPLICATION: US/08/894,356D

TIME: 08:45:08

Input Set : A:\72370308.app

Output Set: N:\CRF4\10112005\H894356D.raw

3 <110> APPLICANT: ASHIKARI, TOSHIHIKO  
 4 TANAKA, YOSHIKAZU  
 5 FUJIWARA, HIROYUKI  
 6 NAKAO, MASAHIRO  
 7 FUKUI, YUKO  
 8 SAKAKIBARA, KEIKO  
 9 MIZUTANI, MASAOKO  
 10 KUSUMI, TAKAOKI  
 12 <120> TITLE OF INVENTION: A GENE ENCODING A PROTEIN HAVING ACYL GROUP TRANSFER  
 13 ACTIVITY  
 15 <130> FILE REFERENCE: 47237-0308/US  
 17 <140> CURRENT APPLICATION NUMBER: 08/894,356D  
 18 <141> CURRENT FILING DATE: 1997-08-18  
 20 <150> PRIOR APPLICATION NUMBER: JP 7-67159  
 21 <151> PRIOR FILING DATE: 1995-02-17  
 23 <150> PRIOR APPLICATION NUMBER: JP 7-196915  
 24 <151> PRIOR FILING DATE: 1995-06-29  
 26 <150> PRIOR APPLICATION NUMBER: JP 8-46534  
 27 <151> PRIOR FILING DATE: 1996-01-30  
 29 <150> PRIOR APPLICATION NUMBER: PCT/JP96/00348  
 30 <151> PRIOR FILING DATE: 1996-02-16  
 32 <160> NUMBER OF SEQ ID NOS: 40  
 34 <170> SOFTWARE: PatentIn Ver. 3.3  
 36 <210> SEQ ID NO: 1  
 37 <211> LENGTH: 1703  
 38 <212> TYPE: DNA  
 39 <213> ORGANISM: Gentiana triflora  
 41 <220> FEATURE:  
 42 <221> NAME/KEY: CDS  
 43 <222> LOCATION: (6)..(1412)  
 45 <400> SEQUENCE: 1  
 46 tcatt atg gag caa atc caa atg gtg aag gtt ctt gaa aaa tgc caa gtt 50  
 47 Met Glu Gln Ile Gln Met Val Lys Val Leu Glu Lys Cys Gln Val  
 48 1 5 10 15  
 50 aca cca cca tct gac aca aca gat gtc gag tta tcg cta ccg gta aca 98  
 51 Thr Pro Pro Ser Asp Thr Thr Asp Val Glu Leu Ser Leu Pro Val Thr  
 52 20 25 30  
 54 ttc ttc gat atc ccc tgg ttg cac ttg aat aag atg cag tcc ctt ctg 146  
 55 Phe Phe Asp Ile Pro Trp Leu His Leu Asn Lys Met Gln Ser Leu Leu  
 56 35 40 45  
 58 ttt tac gac ttt ccg tac cca aga aca cat ttc ttg gac act gtt atc 194  
 59 Phe Tyr Asp Phe Pro Tyr Pro Arg Thr His Phe Leu Asp Thr Val Ile  
 60 50 55 60

## RAW SEQUENCE LISTING

DATE: 10/11/2005

PATENT APPLICATION: US/08/894,356D

TIME: 08:45:08

Input Set : A:\72370308.app

Output Set: N:\CRF4\10112005\H894356D.raw

62	cct	aat	ctt	aag	gcc	tct	ttg	tct	ctc	act	cta	aaa	cac	tac	gtt	ccg	242
63	Pro	Asn	Leu	Lys	Ala	Ser	Leu	Ser	Leu	Thr	Leu	Lys	His	Tyr	Val	Pro	
64		65					70					75					
66	ctt	agc	gga	aat	ttg	ttg	atg	ccg	atc	aaa	tcg	ggc	gaa	atg	ccg	aag	290
67	Leu	Ser	Gly	Asn	Leu	Leu	Met	Pro	Ile	Lys	Ser	Gly	Glu	Met	Pro	Lys	
68	80					85				90						95	
70	ttt	cag	tac	tcc	cgt	gat	gag	ggc	gac	tcg	ata	act	ttg	atc	gtt	gcg	338
71	Phe	Gln	Tyr	Ser	Arg	Asp	Glu	Gly	Asp	Ser	Ile	Thr	Leu	Ile	Val	Ala	
72				100					105						110		
74	gag	tct	gac	cag	gat	ttt	gac	tac	ctt	aaa	ggt	cat	caa	ctg	gta	gat	386
75	Glu	Ser	Asp	Gln	Asp	Phe	Asp	Tyr	Leu	Lys	Gly	His	Gln	Leu	Val	Asp	
76				115					120						125		
78	tcc	aat	gat	ttg	cat	ggc	ctt	ttt	tat	gtt	atg	cca	cgg	gtt	ata	agg	434
79	Ser	Asn	Asp	Leu	His	Gly	Leu	Phe	Tyr	Val	Met	Pro	Arg	Val	Ile	Arg	
80			130						135					140			
82	acc	atg	caa	gac	tat	aaa	gtg	atc	ccg	ctc	gta	gcc	gtg	caa	gta	acc	482
83	Thr	Met	Gln	Asp	Tyr	Lys	Val	Ile	Pro	Leu	Val	Ala	Val	Gln	Val	Thr	
84		145					150						155				
86	gtt	ttt	cct	aac	cgt	ggc	ata	gcc	gtg	gct	ctg	acg	gca	cat	cat	tca	530
87	Val	Phe	Pro	Asn	Arg	Gly	Ile	Ala	Val	Ala	Leu	Thr	Ala	His	His	Ser	
88	160					165					170					175	
90	att	gca	gat	gct	aaa	agt	ttt	gta	atg	ttc	atc	aat	gct	tggt	gcc	tat	578
91	Ile	Ala	Asp	Ala	Lys	Ser	Phe	Val	Met	Phe	Ile	Asn	Ala	Trp	Ala	Tyr	
92				180						185					190		
94	att	aac	aaa	ttt	ggg	aaa	gac	gcg	gac	ttg	ttg	tcc	gcg	aat	ctt	ctt	626
95	Ile	Asn	Lys	Phe	Gly	Lys	Asp	Ala	Asp	Leu	Leu	Ser	Ala	Asn	Leu	Leu	
96			195						200						205		
98	cca	tct	ttc	gat	aga	tcg	ata	atc	aaa	gat	ctg	tat	ggc	cta	gag	gaa	674
99	Pro	Ser	Phe	Asp	Arg	Ser	Ile	Ile	Lys	Asp	Leu	Tyr	Gly	Leu	Glu	Glu	
100			210						215					220			
102	aca	ttt	tggt	aac	gaa	atg	caa	gat	ggt	ctt	gaa	atg	ttc	tct	aga	ttt	722
103	Thr	Phe	Trp	Asn	Glu	Met	Gln	Asp	Val	Leu	Glu	Met	Phe	Ser	Arg	Phe	
104		225						230					235				
106	gga	agc	aaa	ccc	cct	cga	ttc	aac	aag	gta	cga	gct	aca	tat	gtc	ctc	770
107	Gly	Ser	Lys	Pro	Pro	Arg	Phe	Asn	Lys	Val	Arg	Ala	Thr	Tyr	Val	Leu	
108	240					245					250				255		
110	tcc	ctt	gct	gaa	atc	cag	aag	cta	aag	aac	aaa	gta	ctg	aat	ctc	aga	818
111	Ser	Leu	Ala	Glu	Ile	Gln	Lys	Leu	Lys	Asn	Lys	Val	Leu	Asn	Leu	Arg	
112				260						265					270		
114	gga	tcc	gaa	ccg	aca	ata	cgt	gta	acg	acg	ttc	aca	atg	acg	tgt	gga	866
115	Gly	Ser	Glu	Pro	Thr	Ile	Arg	Val	Thr	Thr	Phe	Thr	Met	Thr	Cys	Gly	
116			275						280					285			
118	tac	gta	tggt	aca	tgc	atg	gtc	aaa	tca	aaa	gat	gac	gtc	gta	tca	gag	914
119	Tyr	Val	Trp	Thr	Cys	Met	Val	Lys	Ser	Lys	Asp	Asp	Val	Val	Ser	Glu	
120			290						295				300				
122	gaa	tca	tcg	aac	gac	gaa	aat	gag	ctc	gag	tac	ttc	agt	ttt	aca	gcg	962
123	Glu	Ser	Ser	Asn	Asp	Glu	Asn	Glu	Leu	Glu	Tyr	Phe	Ser	Phe	Thr	Ala	
124		305					310					315					
126	gat	tgc	cga	gga	ctt	ctg	acg	ccc	ccg	tgt	ccg	cct	aac	tac	ttt	ggc	1010

## RAW SEQUENCE LISTING

DATE: 10/11/2005

PATENT APPLICATION: US/08/894,356D

TIME: 08:45:08

Input Set : A:\72370308.app

Output Set: N:\CRF4\10112005\H894356D.raw

```

127 Asp Cys Arg Gly Leu Leu Thr Pro Pro Cys Pro Pro Asn Tyr Phe Gly
128 320 325 330
130 aac tgt ctt gcg tca tgc gtt gca aaa gca aca cat aaa gag tta gtt 1058
131 Asn Cys Leu Ala Ser Cys Val Ala Lys Ala Thr His Lys Glu Leu Val
132 340 345 350
134 ggg gat aaa ggg ctt ctt gtt gca gtt gca gct att gga gaa gcc att 1106
135 Gly Asp Lys Gly Leu Leu Val Ala Val Ala Ala Ile Gly Glu Ala Ile
136 355 360 365
138 gaa aag agg ttg cac aac gaa aaa ggc gtt ctt gca gat gca aaa act 1154
139 Glu Lys Arg Leu His Asn Glu Lys Gly Val Leu Ala Asp Ala Lys Thr
140 370 375 380
142 tgg tta tcg gaa tct aat gga atc cct tca aaa aga ttt ctc ggg att 1202
143 Trp Leu Ser Glu Ser Asn Gly Ile Pro Ser Lys Arg Phe Leu Gly Ile
144 385 390 395
146 acc gga tcg cct aag ttc gat tcg tat ggt gta gat ttt gga tgg gga 1250
147 Thr Gly Ser Pro Lys Phe Asp Ser Tyr Gly Val Asp Phe Gly Trp Gly
148 400 405 410 415
150 aag cct gca aaa ttt gac att acc tct gtt gat tat gca gaa ttg att 1298
151 Lys Pro Ala Lys Phe Asp Ile Thr Ser Val Asp Tyr Ala Glu Leu Ile
152 420 425 430
154 tat gtg att cag tcc agg gat ttt gaa aaa ggt gtg gag att gga gta 1346
155 Tyr Val Ile Gln Ser Arg Asp Phe Glu Lys Gly Val Glu Ile Gly Val
156 435 440 445
158 tca ttg cct aag att cat atg gat gca ttt gca aaa atc ttt gaa gaa 1394
159 Ser Leu Pro Lys Ile His Met Asp Ala Phe Ala Lys Ile Phe Glu Glu
160 450 455 460
162 ggc ttt tgc tct ttg tca tagtctcttt aatagaacca tatttgctgc 1442
163 Gly Phe Cys Ser Leu Ser
164 465
166 aataaagtac caagtccttt agtaacacta caccaaacc tactttcgag gcgggaacac 1502
168 cacaacgagg ttcaatcact agaaggttgt acttcataaa ttccagaggt cgaatataca 1562
170 ccgttgctct ctgaaaagtt gaacctcaca cctgacatgg tgttacgata ggtattgtat 1622
172 aatgccatta tatacttcca taaagtatcc tatgcaatag agaacatggt atgtgttaaa 1682
174 aaaaaaaaaa a 1703
178 <210> SEQ ID NO: 2
179 <211> LENGTH: 1622
180 <212> TYPE: DNA
181 <213> ORGANISM: Gentiana triflora
183 <220> FEATURE:
184 <221> NAME/KEY: CDS
185 <222> LOCATION: (35)..(1471)
187 <400> SEQUENCE: 2
188 gaaccattga atccaattaa tctgatttat taag atg gca gga aat tcc gag gat 55
189 Met Ala Gly Asn Ser Glu Asp
190 1 5
192 atc aaa gtt ctt gag aaa tgc cgt gtt gcg cca cca ccg gac gcc gtc 103
193 Ile Lys Val Leu Glu Lys Cys Arg Val Ala Pro Pro Pro Asp Ala Val
194 10 15 20
196 gcc gag ttt aca gtc cca ctg tcg ttt ttc gac atg cga tgg ttg atc 151

```

## RAW SEQUENCE LISTING

DATE: 10/11/2005

PATENT APPLICATION: US/08/894,356D

TIME: 08:45:08

Input Set : A:\72370308.app

Output Set: N:\CRF4\10112005\H894356D.raw

197	Ala	Glu	Phe	Thr	Val	Pro	Leu	Ser	Phe	Phe	Asp	Met	Arg	Trp	Leu	Ile	
198		25					30				35						
200	tct	gat	gca	gaa	cac	cat	ctg	cat	ttc	tac	aga	ttc	cgc	cat	cct	tgt	199
201	Ser	Asp	Ala	Glu	His	His	Leu	His	Phe	Tyr	Arg	Phe	Arg	His	Pro	Cys	
202	40					45					50					55	
204	ccc	aac	tct	aaa	ttt	atc	att	tca	tcc	att	aaa	tcg	tcc	ctt	tcc	ctt	247
205	Pro	Asn	Ser	Lys	Phe	Ile	Ile	Ser	Ser	Ile	Lys	Ser	Ser	Leu	Ser	Leu	
206					60					65					70		
208	gtt	ctc	aaa	cac	ttt	ctt	ccg	tta	gcc	ggg	aat	ttg	att	tgg	ccg	gta	295
209	Val	Leu	Lys	His	Phe	Leu	Pro	Leu	Ala	Gly	Asn	Leu	Ile	Trp	Pro	Val	
210				75					80					85			
212	gat	tcc	tcc	gat	aga	atg	ccg	gag	ttg	cgt	tac	aag	aaa	ggg	gac	tcc	343
213	Asp	Ser	Ser	Asp	Arg	Met	Pro	Glu	Leu	Arg	Tyr	Lys	Lys	Gly	Asp	Ser	
214			90					95				100					
216	gtt	tct	tta	aca	att	gca	gaa	tcg	agc	atg	gat	ttt	gat	tat	ctc	gcc	391
217	Val	Ser	Leu	Thr	Ile	Ala	Glu	Ser	Ser	Met	Asp	Phe	Asp	Tyr	Leu	Ala	
218		105					110					115					
220	gga	gat	cat	cag	agg	gat	tct	tat	aaa	ttc	aac	gat	ttg	att	ccg	cag	439
221	Gly	Asp	His	Gln	Arg	Asp	Ser	Tyr	Lys	Phe	Asn	Asp	Leu	Ile	Pro	Gln	
222	120					125					130				135		
224	ctg	cca	gaa	ccg	att	gta	acc	tcc	ggc	gac	gaa	gta	tta	cca	ctt	ttt	487
225	Leu	Pro	Glu	Pro	Ile	Val	Thr	Ser	Gly	Asp	Glu	Val	Leu	Pro	Leu	Phe	
226				140						145					150		
228	gct	tta	cag	gtg	acg	gtg	ttc	tcc	aac	acc	ggt	ata	tgc	att	gga	cgc	535
229	Ala	Leu	Gln	Val	Thr	Val	Phe	Ser	Asn	Thr	Gly	Ile	Cys	Ile	Gly	Arg	
230			155						160					165			
232	aat	ctt	cat	caa	gtt	ctt	ggt	gat	gcc	agt	tct	ttt	ctg	cat	ttt	aat	583
233	Asn	Leu	His	Gln	Val	Leu	Gly	Asp	Ala	Ser	Ser	Phe	Leu	His	Phe	Asn	
234			170					175				180					
236	aaa	tta	tgg	gtt	ttg	gtt	gac	aaa	tcc	aat	gga	gat	tca	tta	aag	ttc	631
237	Lys	Leu	Trp	Val	Leu	Val	Asp	Lys	Ser	Asn	Gly	Asp	Ser	Leu	Lys	Phe	
238		185					190					195					
240	ctt	cca	ctt	tct	tct	cta	cct	atg	tac	gac	aga	tct	gtg	gtg	caa	gat	679
241	Leu	Pro	Leu	Ser	Ser	Leu	Pro	Met	Tyr	Asp	Arg	Ser	Val	Val	Gln	Asp	
242	200					205					210				215		
244	cca	ttt	cat	att	cgt	cga	aaa	atc	tac	aat	gaa	aga	aaa	ctg	ctc	aaa	727
245	Pro	Phe	His	Ile	Arg	Arg	Lys	Ile	Tyr	Asn	Glu	Arg	Lys	Leu	Leu	Lys	
246				220						225				230			
248	tct	cag	ggc	aca	cct	act	gtt	cta	aat	cca	gca	att	tct	aaa	gat	gaa	775
249	Ser	Gln	Gly	Thr	Pro	Thr	Val	Leu	Asn	Pro	Ala	Ile	Ser	Lys	Asp	Glu	
250			235						240					245			
252	gtt	cga	gcc	acc	ttc	atc	cta	cac	cct	att	gat	atc	atg	aag	ctc	aag	823
253	Val	Arg	Ala	Thr	Phe	Ile	Leu	His	Pro	Ile	Asp	Ile	Met	Lys	Leu	Lys	
254			250					255					260				
256	aaa	ttc	att	tcg	tca	aaa	aat	cgc	aac	tta	acc	ggt	agt	agt	aat	tat	871
257	Lys	Phe	Ile	Ser	Ser	Lys	Asn	Arg	Asn	Leu	Thr	Gly	Ser	Ser	Asn	Tyr	
258		265					270					275					
260	aat	ctg	tca	act	ttc	acg	gtg	aca	tct	gca	ctg	atc	tgg	aca	tgc	ttg	919
261	Asn	Leu	Ser	Thr	Phe	Thr	Val	Thr	Ser	Ala	Leu	Ile	Trp	Thr	Cys	Leu	

## RAW SEQUENCE LISTING

DATE: 10/11/2005

PATENT APPLICATION: US/08/894,356D

TIME: 08:45:08

Input Set : A:\72370308.app

Output Set: N:\CRF4\10112005\H894356D.raw

```

262 280          285          290          295
264 tcg aaa tca tta gac acc gtc gta aga gag aag gtg gaa gag gat aaa 967
265 Ser Lys Ser Leu Asp Thr Val Val Arg Glu Lys Val Glu Glu Asp Lys
266          300          305          310
268 cat gca gca aac tta tgt gct ttc atc aac tgc cga caa cgt ttt gct 1015
269 His Ala Ala Asn Leu Cys Ala Phe Ile Asn Cys Arg Gln Arg Phe Ala
270          315          320          325
272 ccg ccg ata cct caa aat tac ttt gga aat tgc ata gtg cct tgt atg 1063
273 Pro Pro Ile Pro Gln Asn Tyr Phe Gly Asn Cys Ile Val Pro Cys Met
274          330          335          340
276 gtg gga tcg act cat gag caa ctt gta gga aat gaa ggg ttg tcg gta 1111
277 Val Gly Ser Thr His Glu Gln Leu Val Gly Asn Glu Gly Leu Ser Val
278          345          350          355
280 gct gca acc gcc atc gga gat gct atc cat aag agg tta cat gac tac 1159
281 Ala Ala Thr Ala Ile Gly Asp Ala Ile His Lys Arg Leu His Asp Tyr
282 360          365          370          375
284 gaa gga att ctg aga gga gat tgg ata tcg ccg ccc cga tca aca tct 1207
285 Glu Gly Ile Leu Arg Gly Asp Trp Ile Ser Pro Pro Arg Ser Thr Ser
286          380          385          390
288 gcg gca cca agg tcg acg ctc att tat gtc gtt gga tcc gca caa cgc 1255
289 Ala Ala Pro Arg Ser Thr Leu Ile Tyr Val Val Gly Ser Ala Gln Arg
290          395          400          405
292 aat gtg cat gat ttt gat gca gat ttt ggt tgg gga aag ctt gaa aag 1303
293 Asn Val His Asp Phe Asp Ala Asp Phe Gly Trp Gly Lys Leu Glu Lys
294          410          415          420
296 cat gaa tct gtt tca act aat cct tcg gca aca cta att ttg atc tct 1351
297 His Glu Ser Val Ser Thr Asn Pro Ser Ala Thr Leu Ile Leu Ile Ser
298          425          430          435
300 cgg tcc aga aga ttt aaa gga gca ctt gag ctt ggc att tct ttg cct 1399
301 Arg Ser Arg Arg Phe Lys Gly Ala Leu Glu Leu Gly Ile Ser Leu Pro
302 440          445          450          455
304 aag aat agg atg gac gca ttt gcc acc att ttt acg aat ttc atc aat 1447
305 Lys Asn Arg Met Asp Ala Phe Ala Thr Ile Phe Thr Asn Phe Ile Asn
306          460          465          470
308 agt ctc cat gtg agg agc cct ttg taagaaaaaa gtggtatcaa tgtataaaaa 1501
309 Ser Leu His Val Arg Ser Pro Leu
310          475
312 agacagacaa gttatgatgc aacaaatggt ttaggagatt acaaattccat gggaagatgt 1561
314 atcaaaactca tctctctata tatatatatt caattgtttt aaaaaaaaaa aaaaaaaaaa 1621
316 a 1622
320 <210> SEQ ID NO: 3
321 <211> LENGTH: 1605
322 <212> TYPE: DNA
323 <213> ORGANISM: Petunia hybrida
325 <220> FEATURE:
326 <221> NAME/KEY: CDS
327 <222> LOCATION: (67)..(1410)
329 <400> SEQUENCE: 3
330 tgtcgacgaa atccatttca tttcctcttc tttcttgttt ttctaatttc gtcattcattg 60

```

RAW SEQUENCE LISTING ERROR SUMMARY  
PATENT APPLICATION: US/08/894,356D

DATE: 10/11/2005  
TIME: 08:45:09

Input Set : A:\72370308.app  
Output Set: N:\CRF4\10112005\H894356D.raw

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:16; N Pos. 18  
Seq#:22; N Pos. 9,15

**VERIFICATION SUMMARY**

DATE: 10/11/2005

PATENT APPLICATION: US/08/894,356D

TIME: 08:45:09

Input Set : A:\72370308.app

Output Set: N:\CRF4\10112005\H894356D.raw

L:992 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:16 after pos.:0

L:1081 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:22 after pos.:0